GENERALIZATIONS OF ŚWIERCZKOWSKI'S LEMMA AND THE ARITY GAP OF FINITE FUNCTIONS

MIGUEL COUCEIRO AND ERKKO LEHTONEN

Abstract. Świerczkowski's lemma—as it is usually formulated—asserts that if $f\colon A^n \to A$ is an operation on a finite set A, $n \geq 4$, and every operation obtained from f by identifying a pair of variables is a projection, then f is a semiprojection. We generalize this lemma in various ways. First, it is extended to B-valued functions on A instead of operations on A and to essentially at most unary functions instead of projections. Then we characterize the arity gap of functions of small arities in terms of quasi-arity, which in turn provides a further generalization of Świerczkowski's lemma. Moreover, we explicitly classify all pseudo-Boolean functions according to their arity gap. Finally, we present a general characterization of the arity gaps of B-valued functions on arbitrary finite sets A.

1. Introduction

Świerczkowski's lemma has fundamental consequences in universal algebra (see, e.g., [9, 13, 14, 16]). It is usually formulated as follows: given an operation $f\colon A^n \to A$, $n \geq 4$, if every operation obtained from f by identifying a pair of variables is a projection, then f is a semiprojection, i.e., there is a $t \in \{1,\ldots,n\}$ such that $f(a_1,\ldots,a_n) = a_t$ whenever $a_i = a_j$ for some $1 \leq i < j \leq n$. The importance of this result in clone theory was made apparent in several studies, in particular in Rosenberg's classification of minimal clones [16]. This classification was extended to clones of multioperations in [13], where Świerczkowski's lemma was accordingly adjusted for multiprojections in the multi-valued case.

Świerczkowski's lemma can be generalized in terms of a quasi-ordering of functions $f\colon A^n \to B$, the so-called simple minor relation [4, 5]: a function f is a simple minor of a function g if f can be obtained from g by permutation of variables, addition or deletion of inessential variables, and identification of variables (see Section 2). With this terminology, in Section 3 we extend Świerczkowski's lemma to the following: if every function obtained from $f\colon A^n \to B$ by identifying a pair of variables is essentially unary and $n = 2$ or $n \geq 4$ (constant and $n \geq 2$, respectively), then there exists an essentially unary function (a constant function, respectively) $g\colon A^n \to B$ such that $f|_{A^n_=} = g|_{A^n_=}$, where

$$A^n_= = \{(a_1,\ldots,a_n) \in A^n : a_i = a_j \text{ for some } 1 \leq i < j \leq n\}.$$

The simple minor relation plays an important role in the equational approach to function class definability. Ekin *et al.* [7] showed that the equationally definable classes of Boolean functions coincide with the initial segments of this quasi-ordering. This result was extended to functions defined on arbitrary, possibly infinite domains in [3]. These results motivated a study [5] of the simple minor relation where the

structure of the corresponding quasi-ordered set was investigated and where the notion of arity gap had useful applications.

The arity gap of a function $f\colon A^n \to B$ which depends on all of its variables, $n \geq 2$, is defined as the minimum decrease in the number of essential variables when variables of f are identified. The arity gap of a function is obviously at least 1, and it can be as large as $|A|$, as shown by the following example due to Salomaa [17]. Assume that $|A| = k \geq 2$, let $\mathbf{a} = (a_1, \ldots, a_k) \in A^k$ be a fixed k-tuple such that $a_i \neq a_j$ whenever $i \neq j$, and let b and c be distinct elements of A. Define the operation $f\colon A^k \to A$ as follows:

$$f(\mathbf{x}) = \begin{cases} b, & \text{if } \mathbf{x} = \mathbf{a}, \\ c, & \text{otherwise.} \end{cases}$$

All k variables are essential in f, and whenever any two variables are identified, the resulting function is constant, having no essential variables. Thus the arity gap of f is k.

Salomaa [17] showed that the arity gap of any Boolean function is at most 2. This result was extended to functions defined on arbitrary finite domains by Willard [20], who showed that the same upper bound holds for the arity gap of any function $f\colon A^n \to B$, provided that $n > |A|$. In fact, he showed that if the arity gap of such a function f is 2, then f is totally symmetric. In Section 4 we consider the arity gap of functions $f\colon A^n \to B$ of small arities, namely, where $n \leq |A|$. We give a characterization of such functions by introducing the notion of quasi-arity, which in turn leads to a further generalization of Świerczkowski's lemma.

In [4] we strengthened Salomaa's [17] result on the upper bound for the arity gap of Boolean functions by completely classifying the Boolean functions according to their arity gap. By making use of tools provided by Berman and Kisielewicz [1] and Willard [20], we obtain in Section 5 a similar explicit classification of all pseudo-Boolean functions, i.e., functions $f\colon \{0,1\}^n \to B$, where B is an arbitrary set. In Section 6, we present a general characterization of finite functions according to their arity gap which is given in terms of quasi-arity.

2. Simple variable substitutions

Throughout this paper, let A be an arbitrary finite set with $|A| = k \geq 2$ elements, and let B be an arbitrary set with at least two elements. A *B-valued function (of several variables) on A* is a mapping $f\colon A^n \to B$ for some positive integer n, called the *arity* of f. *A*-valued functions on A are called *operations on A*. Operations on $\{0,1\}$ are called *Boolean functions*. For an arbitrary B, we refer to B-valued functions on $\{0,1\}$ as *pseudo-Boolean functions*.

For each positive integer n, the *n*-ary *projections* $(a_1, \ldots, a_n) \mapsto a_i$, $1 \leq i \leq n$, are also called *variables* and denoted by $x_i^{(n)}$, or simply by x_i when the arity is clear from the context. We say that the i-th variable x_i is *essential* in f, or f *depends* on x_i, if there are elements $a_1, \ldots, a_n, b \in A$ such that

$$f(a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n) \neq f(a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n).$$

In this case, the pair $((a_1, \ldots, a_{i-1}, a_i, a_{i+1}, \ldots, a_n), (a_1, \ldots, a_{i-1}, b, a_{i+1}, \ldots, a_n))$ is called a *witness of essentiality* of x_i in f. If a variable is not essential in f, then we say that it is *inessential* in f. The number of essential variables in f is called the *essential arity* of f, and it is denoted by $\operatorname{ess} f$. If $\operatorname{ess} f = m$, we say that f

is *essentially m-ary.* Thus the only essentially nullary functions are the constant functions.

We extend the notion of essential variable to partial functions $f\colon S \to B$, where $S \subseteq A^n$. The definition is in fact the same as the one for total functions $A^n \to B$, but now the witnesses of essentiality must be in S^2. In other words, we say that the i-th variable x_i is *essential* in $f\colon S \to B$, where $S \subseteq A^n$, or f *depends* on x_i, if there is a pair

$$((a_1,\ldots,a_{i-1},a_i,a_{i+1},\ldots,a_n),(a_1,\ldots,a_{i-1},b,a_{i+1},\ldots,a_n)) \in S^2,$$

called a *witness of essentiality* of x_i in f, such that

$$f(a_1,\ldots,a_{i-1},a_i,a_{i+1},\ldots,a_n) \neq f(a_1,\ldots,a_{i-1},b,a_{i+1},\ldots,a_n).$$

With no risk of ambiguity, the notion of essential arity is defined for partial functions as in the case of total functions.

Suppose that $f\colon S \to B$ is a partial function with $S \subseteq A^n$, and $\operatorname{ess} f = m$. Let us suppose without loss of generality (renaming the variables if necessary) that the essential variables are $x_1,\ldots,x_m$. Consider the statement

$$(1) \qquad \exists h\colon A^m \to B : \forall (a_1,\ldots,a_n) \in S : f(a_1,\ldots,a_n) = h(a_1,\ldots,a_m).$$

Statement (1) obviously holds for total functions, i.e., when $S = A^n$. It is not true for partial functions in general. (For a counterexample, assume that $|A| \geq 2$, let $S = \{(a,a,\ldots,a) : a \in A\} \subseteq A^n$, and define $f\colon S \to A$ by $f(a,a,\ldots,a) = a$ for all $(a,a,\ldots,a) \in S$. All variables of f are inessential, yet f is not a constant function.) However, (1) is true for partial functions whose domain has a special shape, as described below.

For $n \geq 2$, define the set

$$A_{\underline{=}}^n = \{(a_1,\ldots,a_n) \in A^n : a_i = a_j \text{ for some } i \neq j\}.$$

We define $A_{\underline{=}}^1 = A$. Note that if $n > |A|$, then $A_{\underline{=}}^n = A^n$.

Lemma 1. *Let $f\colon A_{\underline{=}}^n \to B$ be a partial function, $n \neq 2$, and $\operatorname{ess} f = m$. Suppose that the essential variables of f are $x_1,\ldots,x_m$. Then there exists a total function $h\colon A^m \to B$ such that for all $(a_1,\ldots,a_n) \in A_{\underline{=}}^n$, $f(a_1,\ldots,a_n) = h(a_1,\ldots,a_m)$.*

Proof. The case $n = 1$ being trivial, we may assume that $n \geq 3$. If $m = n$, then every extension of f into a total function satisfies the required condition. Thus, we can assume that $m < n$. Let $h(a_1,\ldots,a_m) = f(a_1,\ldots,a_m,a_m,\ldots,a_m)$ for all $(a_1,\ldots,a_n) \in A^n$. Then for any $(a_1,\ldots,a_n) \in A_{\underline{=}}^n$ we have

$f(a_1,\ldots,a_m,a_{m+1},a_{m+2},a_{m+3},\ldots,a_{n-1},a_n)$

$= f(a_1,\ldots,a_m,a_m,a_{m+2},a_{m+3},\ldots,a_{n-1},a_n)$ (since x_{m+1} is inessential in f)

$= f(a_1,\ldots,a_m,a_m,a_m,a_{m+3},\ldots,a_{n-1},a_n)$ (since x_{m+2} is inessential in f)

$\vdots$

$= f(a_1,\ldots,a_m,a_m,a_m,a_m,\ldots,a_m,a_m,a_n)$

$= f(a_1,\ldots,a_m,a_m,a_m,a_m,\ldots,a_m,a_m,a_m)$ (since x_n is inessential in f)

$= h(a_1,\ldots,a_m).$

(Note that the n-tuples we used all belong to $A_{\underline{=}}^n$.) $\square$

The *composition* of $f: B^n \to C$ with $g_1, \dots, g_n: A^m \to B$ is the function $f(g_1, \dots, g_n): A^m \to C$ defined by

$$f(g_1, \dots, g_n)(\mathbf{a}) = f(g_1(\mathbf{a}), \dots, g_n(\mathbf{a}))$$

for all $\mathbf{a} \in A^m$.

We say that a function $f: A^n \to B$ is obtained from $g: A^m \to B$ by *simple variable substitution,* or f is a *simple minor* of g, if there is a mapping $\sigma: \{1, \dots, m\} \to \{1, \dots, n\}$ such that

$$f = g(x^{(n)}_{\sigma(1)}, \dots, x^{(n)}_{\sigma(m)}).$$

If σ is not injective, then we speak of *identification of variables.* If σ is not surjective, then we speak of *addition of inessential variables.* If σ is a bijection, then we speak of *permutation of variables.* Observe that each $x^{(n)}_{\sigma(i)}$ is simply an n-ary projection, and thus we have that f is a simple minor of g if and only if

$$\{f(\pi_1, \dots, \pi_n) : \pi_1, \dots, \pi_n \text{ are projections of the same arity}\} \subseteq$$

$$\{g(\rho_1, \dots, \rho_m) : \rho_1, \dots, \rho_m \text{ are projections of the same arity}\}.$$

From this observation it follows that the simple minor relation constitutes a quasi-order $\leq$ on the set of all B-valued functions of several variables on A which is given by the following rule: $f \leq g$ if and only if f is obtained from g by simple variable substitution. If $f \leq g$ and $g \leq f$, we say that f and g are *equivalent,* denoted $f \equiv g$. If $f \leq g$ but $g \not\leq f$, we denote $f < g$. It can be easily observed that if $f \leq g$ then $\operatorname{ess} f \leq \operatorname{ess} g$, with equality if and only if $f \equiv g$. For background, extensions and variants of the simple minor relation, see, e.g., [2, 5, 8, 10, 11, 12, 19, 21].

3. Quasi-arity and a generalization of Świerczkowski's lemma

In this section, we extend Świerczkowski's lemma to B-valued functions on A. To this extent, we need to introduce some terminology.

Let $f: A^n \to B$, where $n \geq 2$. For indices $i, j \in \{1, \dots, n\}$, $i \neq j$, the function $f_{i \leftarrow j}: A^n \to B$ obtained from $f: A^n \to B$ by the simple variable substitution

$$f_{i \leftarrow j} = f(x^{(n)}_1, \dots, x^{(n)}_{i-1}, x^{(n)}_j, x^{(n)}_{i+1}, \dots, x^{(n)}_n)$$

is called a *variable identification minor* of f, obtained by identifying x_i with x_j.

The *diagonal function* of $f: A^n \to B$ is the mapping $\Delta_f: A \to B$ defined by $\Delta_f(a) = f(a, a, \dots, a)$ for all $a \in A$. Equivalently, in terms of functional composition, $\Delta_f = f(x^{(1)}_1, \dots, x^{(1)}_1)$.

The proofs of the following two lemmas are straightforward and are left to the reader.

Lemma 2. $\Delta_f = \Delta_{f_{i \leftarrow j}}$ *for all* $i \neq j$.

Lemma 3. *A function* $f: A^n \to B$ *is essentially at most unary if and only if* $f = \Delta_f(x^{(n)}_i)$ *for some* $1 \leq i \leq n$.

Let $f: A^n \to B$. Any function $g: A^n \to B$ satisfying $f|_{A^n_{\underline{\underline{}}}} = g|_{A^n_{\underline{\underline{}}}}$ is called a *support* of f. The *quasi-arity* of f, denoted $\operatorname{qa} f$, is defined as the minimum of the essential arities of the supports of f, i.e., $\operatorname{qa} f = \min \operatorname{ess} g$, where g ranges over the set of all supports of f. If $\operatorname{qa} f = m$, we say that f is *quasi-m-ary.* We call f a *semiprojection,* if there exists a projection that is a support of f, in other

words, if there is a $t \in \{1, \ldots, n\}$ such that $f(a_1, \ldots, a_n) = a_t$ whenever $a_i = a_j$ for some $1 \leq i < j \leq n$. Note that according to our definition, all projections are semiprojections. Some authors do not consider projections as semiprojections (e.g., [6, 15]).

Remark 1.   (i) If $n > |A|$, then quasi-m-ary means the same as essentially m-ary, in particular, quasi-nullary and quasi-unary mean the same as constant and essentially unary, respectively.

(ii) The quasi-arity of $f \colon A^2 \to B$ is either 0 or 1, depending on whether the diagonal function Δ_f is constant or nonconstant, respectively. Furthermore, the two possible variable identification minors $f_{1 \leftarrow 2}$ and $f_{2 \leftarrow 1}$ are equivalent to Δ_f.

(iii) Among all supports of a quasi-nullary function, there is exactly one which is constant. If $f \colon A^2 \to B$ is quasi-unary, then it has exactly two essentially unary supports. For $n \geq 3$, if $f \colon A^n \to B$ is quasi-unary, then among all supports of f there is exactly one which is essentially unary.

The following lemma establishes the connection between the quasi-arity of f and the essential arity of the restriction $f|_{A^n_{\underline{\equiv}}}$.

Lemma 4. *For every function $f \colon A^n \to B$, $n \neq 2$, we have* $\operatorname{qa} f = \operatorname{ess} f|_{A^n_{\underline{\equiv}}}$.

Proof. The case when $n = 1$ being trivial, we may assume that $n \geq 3$. Since any witness of essentiality of an essential variable x_i in $f|_{A^n_{\underline{\equiv}}}$ is also a witness of essentiality of x_i in any support of f, every essential variable of $f|_{A^n_{\underline{\equiv}}}$ must be essential in every support of f, and therefore we have $\operatorname{qa} f \geq \operatorname{ess} f|_{A^n_{\underline{\equiv}}}$.

Let $\operatorname{ess} f|_{A^n_{\underline{\equiv}}} = m$, and assume without loss of generality that the essential variables of $f|_{A^n_{\underline{\equiv}}}$ are $x_1, \ldots, x_m$. By Lemma 1, there is a function $h \colon A^m \to B$ such that for all $(a_1, \ldots, a_n) \in A^n_{\underline{\equiv}}$, $f|_{A^n_{\underline{\equiv}}}(a_1, \ldots, a_n) = h(a_1, \ldots, a_m)$. By introducing $n - m$ inessential variables, we obtain the function $h' \colon A^n \to B$ given by $h'(a_1, \ldots, a_n) = h(a_1, \ldots, a_m)$ for all $(a_1, \ldots, a_n) \in A^n$, and it is clear that $\operatorname{ess} h' = \operatorname{ess} h$. Then h' is a support of f, and since h has arity m, we have $\operatorname{qa} f \leq \operatorname{ess} h' = \operatorname{ess} h \leq m$. $\square$

Note that Lemma 4 does not hold for $n = 2$. By Remark 1(ii), the quasi-arity of a binary fuction $f \colon A^2 \to B$ is either 0 or 1, but $\operatorname{ess} f|_{A^2_{\underline{\equiv}}} = 0$.

Lemma 5. *If a quasi-m-ary function $f \colon A^n \to B$ has an inessential variable, then f is essentially m-ary.*

Proof. The statement is easily seen to hold when $n < 3$, so we may assume that $n \geq 3$. Let $\operatorname{qa} f = m$. By Lemma 4, $\operatorname{ess} f|_{A^n_{\underline{\equiv}}} = m$. Assume without loss of generality that the essential variables of $f|_{A^n_{\underline{\equiv}}}$ are $x_1, \ldots, x_m$. Every essential variable of $f|_{A^n_{\underline{\equiv}}}$ is obviously essential in f, so $\operatorname{ess} f \geq m$. Lemma 1 implies that there is a function $h \colon A^m \to B$ such that for all $(a_1, \ldots, a_n) \in A^n_{\underline{\equiv}}$ we have $f(a_1, \ldots, a_n) = h(a_1, \ldots, a_m)$. If a variable, say x_n, is inessential in f, then for any $(a_1, \ldots, a_n) \in A^n$ we have $f(a_1, \ldots, a_{n-1}, a_n) = f(a_1, \ldots, a_{n-1}, a_{n-1}) = h(a_1, \ldots, a_m)$. This shows that $\operatorname{ess} f \leq m$. $\square$

Świerczkowski [18, statement (β) in Section 2] proves the following lemma about partitions of finite sets. Here, for any partitions δ, δ' of a set S, we write $\delta < \delta'$ if, for each block D of δ, there is a block D' of δ' which contains D.

Lemma 6. *If we have a fixed number $n \geq 3$, S is a finite set, and to every partition δ of S in not more than n disjoint subsets corresponds a set $\varphi\delta$ of that partition so that $\delta < \delta'$ implies $\varphi\delta < \varphi\delta'$, then the intersection of all $\varphi\delta$ is nonempty.*

Lemma 6 forms the base of the theorem which is usually named as Świerczkowski's lemma: Let f be a function of arity at least 4, such that every simple minor of f is a projection. If δ is a nontrivial partition (different from equality) of the set of variables, then identifying the variables belonging to the same block of δ, we obtain a simple minor f_δ of f. Let $\varphi\delta$ denote that block of δ which contains the variable to which f_δ projects. Then applying Lemma 6 we find that the intersection of the sets $\varphi\delta$ is not empty; thus f_δ is always the same projection, i.e., f is a semiprojection.

Theorem 7. *Let $f \colon A^n \to B$.*

  (i) *For $n \geq 2$, all variable identification minors of f are constant functions if and only if f is quasi-nullary.*

  (ii) *For $n = 2$ or $n \geq 4$, all variable identification minors of f are essentially unary if and only if f is quasi-unary.*

Furthermore, in (i) and in (ii), provided that $n \geq 4$, the variable identification minors of f are equivalent to the unique essentially at most unary support of f.

Proof. By Lemmas 2 and 3, it is not possible that f has both constant functions and essentially unary functions as minors. Thus, in light of Remark 1(ii), part (ii) for $n = 2$ will follow from part (i).

(i) It is clear by definition that all minors of a quasi-nullary function are constant. For the converse implication, assume that all minors of f are constant. Since $\Delta_f = \Delta_{f_{i \leftarrow j}}$ for all $i \neq j$ by Lemma 2, they must in fact be constant functions taking on the same value, say $c \in B$. Thus, $f_{i \leftarrow j}(\mathbf{a}) = c$ for all $\mathbf{a} \in A^n$ and for all $i \neq j$, so $f(\mathbf{a}) = c$ for all $\mathbf{a} \in A^n_=$. Hence $f|_{A^n_=}$ is a constant map, and so f is quasi-nullary.

(ii) It is again clear by definition that all minors of a quasi-unary function are essentially unary. For the converse implication, assume that $n \geq 4$ and all minors of f are essentially unary. For a nontrivial partition δ of the set of variables of f, denote by f_δ the simple minor of f that is obtained by identifying the variables belonging to the same block. Let $\varphi\delta$ be that block of δ which contains the only essential variable of f_δ. Then applying Lemma 6 we find that the intersection of the sets $\varphi\delta$ is not empty; thus f_δ always depends on the same essential variable, i.e., f is quasi-unary.

The last claim follows from Lemmas 2 and 3 and Remark 1(iii). This completes the proof of Theorem 7. $\qquad\square$

For $A = B$, Theorem 7 restricted to semiprojections entails the well-known formulation of Świerczkowski's lemma (see, e.g., [6, 15]).

Lemma 8 (Świerczkowski's lemma)**.** *Let f be an n-ary operation on A and $n \geq 4$. Then f is a semiprojection if and only if every variable identification minor of f is a projection.*

4. Arity gap and a further generalization of Świerczkowski's lemma

Recall that simple variable substitution induces a quasi-order on the set of B-valued functions on A, as described in Section 2. For a function $f \colon A^n \to B$ with

at least two essential variables, we denote

$$\mathrm{ess}^< f = \max_{g < f} \mathrm{ess}\, g,$$

and we define the *arity gap* of f by $\mathrm{gap}\, f = \mathrm{ess}\, f - \mathrm{ess}^< f$. It is easily observed that

$$\mathrm{gap}\, f = \min_{i \neq j}(\mathrm{ess}\, f - \mathrm{ess}\, f_{i \leftarrow j}),$$

where i and j range over the set of indices of essential variables of f.

Since the arity gap is defined in terms of essential variables and since every B-valued function on A is equivalent to a function whose variables are all essential, we will assume without loss of generality that the functions $f \colon A^n \to B$ whose arity gap we consider are essentially n-ary and $n \geq 2$.

The following upper bound for the arity gap was established by Willard [20, Lemma 1.2].

Theorem 9. *Suppose $f \colon A^n \to B$ depends on all of its variables. If $n > k$, then* $\mathrm{gap}\, f \leq 2$.

This theorem leaves unsettled the arity gaps of functions with a small number of essential variables, i.e., the case that $2 \leq n \leq k$.

Proposition 10. *Suppose $f \colon A^n \to B$, $2 \leq n \leq k$, depends on all of its variables. If $\mathrm{qa}\, f = m < n$, then $\mathrm{gap}\, f = n - m$.*

Proof. Let g be an essentially m-ary support of f. The variable identification minors of f coincide with those of g, i.e., $f_{i \leftarrow j} = g_{i \leftarrow j}$ for all $i \neq j$, and hence $\mathrm{ess}\, f_{i \leftarrow j} = \mathrm{ess}\, g_{i \leftarrow j} \leq \mathrm{ess}\, g$. Since $m < n$, g has an inessential variable, say x_p, and therefore for any $q \neq p$ we have that $f_{p \leftarrow q} = g_{p \leftarrow q} = g$. Thus, $\mathrm{gap}\, f = n - m$. $\square$

In order to deal with the case that $\mathrm{qa}\, f = n$, we will adapt Willard's proof of Theorem 9 to functions of small essential arity. The idea is that it suffices to consider the restriction of f to $A_{\underline{=}}^n$ and hence the condition $n > k$ can be omitted.

Theorem 11. *Suppose $f \colon A^n \to B$, $n > 3$, depends on all of its variables. If $\mathrm{qa}\, f = n$, then $\mathrm{gap}\, f \leq 2$.*

Proof. Let r be the maximum of the essential arities of all $f_{i \leftarrow j}$, $i \neq j$. Assume on the contrary that $r < n - 2$. We shall find a contradiction.

Claim 1. There exist u, v such that $f_{v \leftarrow u}$ is essentially r-ary and does not depend on x_u.

Claim 1 was proved by Willard (Claim 1 in the proof of Lemma 1.2 in [20]).

Let u and v be as in Claim 1, and assume without loss of generality that $u = n-1$, $v = n$ and the essential variables of $f_{v \leftarrow u}$ are $x_1, \ldots, x_r$. Then

$$f_{v \leftarrow u}(x_1, \ldots, x_n) = f(x_1, \ldots, x_{n-2}, x_{n-1}, x_{n-1}) = h(x_1, \ldots, x_r),$$

where h depends on all of its variables.

Claim 2. For all $(a_1, \ldots, a_n) \in A_{\underline{=}}^n$, $f(a_1, \ldots, a_n) = h(a_1, \ldots, a_r)$.

Claim 2 was proved by Willard (Claim 2 in the proof of Lemma 1.2 in [20]). Claim 2 implies that for all $(a_1, \ldots, a_n) \in A_{\underline{=}}^n$, $f|_{A_{\underline{=}}^n}(a_1, \ldots, a_n) = h(a_1, \ldots, a_r)$, and hence $\mathrm{qa}\, f = \mathrm{ess}\, f|_{A_{\underline{=}}^n} \leq r < n$, which is a contradiction. $\square$

Observe that if $n > k$, then Theorem 11 reduces to Theorem 9, because in this case quasi-arity means the same as essential arity.

From Proposition 10 and Theorem 11 we can now derive the following characterization of functions of arity gap at least 3.

Theorem 12. *Suppose $f\colon A^n \to B$ depends on all of its variables. For $0 \le m \le n - 3$, we have that $\operatorname{gap} f = n - m$ if and only if $\operatorname{qa} f = m$.*

Note that in Theorem 12 the cases $m = 0$ and $m = 1$ correspond to parts (i) and (ii) of Theorem 7, respectively. Thus, Theorem 12 can be viewed as a further generalization of Świerczkowski's lemma.

5. The classification of pseudo-Boolean functions according to their arity gap

In [4], we completely classified all Boolean functions in terms of arity gap. More precisely, we have shown that for a Boolean function $f\colon \{0,1\}^n \to \{0,1\}$ with at least two essential variables, $\operatorname{gap} f = 2$ if and only if f is equivalent to one of the following Boolean functions:

- $x_1 + x_2 + \cdots + x_m + c$ $(2 \le m \le n)$,
- $x_1 x_2 + x_1 + c$,
- $x_1 x_2 + x_1 x_3 + x_2 x_3 + c$,
- $x_1 x_2 + x_1 x_3 + x_2 x_3 + x_1 + x_2 + c$,

where addition and multiplication are done modulo 2 and $c \in \{0,1\}$. Otherwise $\operatorname{gap} f = 1$.

By Theorem 9, the arity gap of a pseudo-Boolean function is either 1 or 2. Like in the case of Boolean functions, this fact asks for a complete classification of pseudo-Boolean functions in terms of arity gap. By making use of tools appearing in [1, 20], we obtain the following characterization.

Theorem 13. *For a pseudo-Boolean function $f\colon \{0,1\}^n \to B$, $n \ge 2$, which depends on all of its variables, $\operatorname{gap} f = 2$ if and only if f satisfies one of the following conditions:*

- *$n = 2$ and f is a nonconstant function satisfying $f(0,0) = f(1,1)$,*
- *$f = g \circ h$, where $g\colon \{0,1\} \to B$ is injective and $h\colon \{0,1\}^n \to \{0,1\}$ is a Boolean function with $\operatorname{gap} h = 2$, as listed above.*

Otherwise $\operatorname{gap} f = 1$.

In order to prove Theorem 13, we need to introduce some terminology and auxiliary results. Let $\mathcal{P}(A)$ denote the power set of A. For each positive integer n, define the function $\operatorname{oddsupp}\colon A^n \to \mathcal{P}(A)$ by

$$\operatorname{oddsupp}(\mathbf{a}) = \{a_i : |\{j \in \{1, \ldots, n\} : a_j = a_i\}| \text{ is odd}\}.$$

A function $f\colon A^n \to B$ is said to be *determined by* $\operatorname{oddsupp}$ if there is a function $f^*\colon \mathcal{P}(A) \to B$ such that $f = f^* \circ \operatorname{oddsupp}$. In [1, 20] it was shown that if $f\colon A^n \to B$ where $n > \max(k, 3)$ and $\operatorname{gap} f = 2$ then f is determined by $\operatorname{oddsupp}$. This result leads to the following auxiliary lemma.

Lemma 14. *Suppose that $f\colon A^n \to B$, where $n > \max(k, 3)$, depends on all of its variables. If the range of f contains more than 2^{k-1} elements, then $\operatorname{gap} f = 1$.*

Proof. For the sake of contradiction, suppose that $\operatorname{gap} f = 2$. Then f is determined by $\operatorname{oddsupp}$, so $f = f^* \circ \operatorname{oddsupp}$ for some $f^*\colon \mathcal{P}(A) \to B$. In fact, the range of

oddsupp contains only subsets of A of even order or only subsets of odd order, depending on the parity of n. The number of subsets of A of even order equals the number of subsets of odd order, and this number is 2^{k-1}. Then the range of f contains at most 2^{k-1} elements, contradicting the hypothesis.   □

Proof of Theorem 13. It is easy to verify that if $f\colon \{0,1\}^n \to B$ satisfies any of the conditions mentioned in the statement of the theorem, then $\operatorname{gap} f = 2$. So we need to show that there are no other functions $f\colon \{0,1\}^n \to B$ with arity gap 2. In fact, we only need to verify the case where the range of f contains at least three elements, because otherwise f is of the form $f = g \circ h$, where $g\colon \{0,1\} \to B$ is injective and $h\colon \{0,1\}^n \to \{0,1\}$ is a Boolean function, and it is clear that in this case $\operatorname{gap} f = \operatorname{gap} g$.

Unary functions cannot have arity gap 2. The case of binary functions is straightforward to verify. If $n > 3$, then Lemma 14 implies that $\operatorname{gap} f = 1$.

If $n = 3$, we have two cases. Consider first the case that $f(0,0,0) = f(1,1,1)$. For $i \neq j$, let g_{ij} be the binary function equivalent to $f_{i \leftarrow j}$. We have that $g_{ij}(0,0) = g_{ij}(1,1)$ and it is easy to see that g_{ij} is either constant or essentially binary. Since $\operatorname{gap} f \leq 2$, the g_{ij} cannot all be constant. Hence, there is an essentially binary g_{ij} for some $i \neq j$, and we conclude that $\operatorname{gap} f = 1$.

Consider then the case that $f(0,0,0) = a \neq b = f(1,1,1)$. Let c be an element in the range of f distinct from both a and b, and let $\mathbf{u} \in A^n$ be such that $f(\mathbf{u}) = c$. Then $\mathbf{u}$ has two coinciding coordinates, say i and j. Let g_{ij} be the binary function equivalent to $f_{i \leftarrow j}$. Then g_{ij} takes on at least three distinct values, namely a, b, c, and it is clear that g_{ij} is essentially binary, and hence $\operatorname{gap} f = 1$.   □

6. General classification of functions according to their arity gap

In the previous section, we presented an explicit classification of pseudo-Boolean functions according to their arity gap. In the general case of functions $f\colon A^n \to B$ where $|A| > 2$, assuming no specific algebraic structure on the underlying set A, such an explicit general description (in terms of representations, e.g., polynomial, DNF, etc.) seems unattainable. Falling short of explicitness, but achieving full generality, in this section we classify finite functions according to their arity gap in terms of quasi-arity and the notion of a function being determined by oddsupp.

Willard's proof of Theorem 2.1 in [20] immediately gives rise to the following generalization, where we have omitted the condition $n > |A|$.

Theorem 15. *Let $f\colon A^n \to B$, $n > 3$, and suppose that $\operatorname{qa} f = n$ and $\operatorname{gap} f = 2$. Then $f|_{A^n_{\underline{=}}}$ is totally symmetric and for all $i \neq j$, $f_{i \leftarrow j}$ depends on all of its variables except x_i and x_j.*

Berman and Kisielewicz's [1] Lemma 2.7 can similarly be generalized, removing the condition $n > |A|$, which together with Theorem 15 and Proposition 10 yields the following result providing a necessary and sufficient condition for a function to have arity gap 2. For $f\colon A^n \to B$, we say that $f|_{A^n_{\underline{=}}}$ *is determined by* oddsupp if $f|_{A^n_{\underline{=}}} = f^* \circ \operatorname{oddsupp}$, where $f^*\colon \mathcal{P}'(A) \to B$ is a nonconstant function and $\operatorname{oddsupp}\colon A^n_{\underline{=}} \to \mathcal{P}'(A)$ is defined as in Section 5, but here $\mathcal{P}'(A)$ denotes the set of odd or even—depending on the parity of n—subsets of A of order at most $n - 2$.

Theorem 16. *Suppose $f\colon A^n \to B$, $n > 3$, depends on all of its variables. Then $\operatorname{gap} f = 2$ if and only if $\operatorname{qa} f = n - 2$ or $\operatorname{qa} f = n$ and $f|_{A^n_{\underline{=}}}$ is determined by oddsupp.*

We now have everything ready to present a complete classification of all functions $f\colon A^n \to B$ according to their arity gap.

Theorem 17. *Suppose that $f\colon A^n \to B$, $n \geq 2$, depends on all of its variables.*
  (i) *For $0 \leq m \leq n - 3$, $\operatorname{gap} f = n - m$ if and only if $\operatorname{qa} f = m$.*
 (ii) *For $n \neq 3$, $\operatorname{gap} f = 2$ if and only if $\operatorname{qa} f = n - 2$ or $\operatorname{qa} f = n$ and $f|_{A^n_{\underline{=}}}$ is determined by oddsupp.*
(iii) *For $n = 3$, $\operatorname{gap} f = 2$ if and only if there is a nonconstant unary function $h\colon A \to B$ and $i_1, i_2, i_3 \in \{0, 1\}$ such that*

$$f(x_1, x_0, x_0) = h(x_{i_1}),$$
$$f(x_0, x_1, x_0) = h(x_{i_2}),$$
$$f(x_0, x_0, x_1) = h(x_{i_3}).$$

(iv) *Otherwise $\operatorname{gap} f = 1$.*

Proof. Statements (i) and (ii) follow from Theorems 12 and 16, respectively.

For (ii), the sufficiency is obvious. In order to prove the necessity, assume that $f\colon A^3 \to B$ has arity gap 2. Then all variable identification minors of f have essential arity at most 1 and at least one of them has essential arity 1. By Lemma 2, all of them have essential arity 1 and are actually of the form $\Delta_f(x_{i_j})$ for some $i_1, i_2, i_3 \in \{1, 2, 3\}$.

Statement (iv) is clear, because (i)–(iii) exhaust all other possibilities.   □

Let f be a ternary function with arity gap 2, and let h, i_1, i_2, i_3 be as in Theorem 17(iii). If $(i_1, i_2, i_3) = (1, 0, 0)$, $(0, 1, 0)$ or $(0, 0, 1)$, then $\operatorname{qa} f = 1$; in the other five cases $\operatorname{qa} f = 3$. If $(i_1, i_2, i_3) = (1, 1, 1)$, then $f|_{A^n_{\underline{\bar{=}}}}$ is determined by oddsupp, otherwise not. Thus $(i_1, i_2, i_3) = (1, 1, 0)$, $(1, 0, 1)$, $(0, 1, 1)$ and $(0, 0, 0)$ give counterexamples to show that statement (ii) of Theorem 17 does not hold when $n = 3$ (and these are the only counterexamples).

Note that in the case $A = B$ and $h = \operatorname{id}_A$ we have that $(i_1, i_2, i_3) = (0, 0, 0)$ if and only if f is a majority operation; $(i_1, i_2, i_3) = (1, 1, 1)$ if and only if f is a minority operation; $(i_1, i_2, i_3) = (1, 0, 0)$, $(0, 1, 0)$ or $(0, 0, 1)$ if and only if f is a semiprojection; and $(i_1, i_2, i_3) = (1, 1, 0)$, $(1, 0, 1)$ or $(0, 1, 1)$ if and only if f is a so-called $\frac{2}{3}$-minority operation.

References

[1] J. Berman, A. Kisielewicz, On the number of operations in a clone, *Proc. Amer. Math. Soc.* **122** (1994) 359–369.
[2] M. Couceiro, On the lattice of equational classes of Boolean functions and its closed intervals, *J. Multiple-Valued Logic Soft Comput.* **18** (2008) 81–104.

[3] M. Couceiro, S. Foldes, Functional equations, constraints, definability of function classes, and functions of Boolean variables, *Acta Cybernet.* **18** (2007) 61–75.

[4] M. Couceiro, E. Lehtonen, On the effect of variable identification on the essential arity of functions on finite sets, *Int. J. Found. Comput. Sci.* **18** (2007) 975–986.

[5] M. Couceiro, M. Pouzet, On a quasi-ordering on Boolean functions, *Theoret. Comput. Sci.* **396** (2008) 71–87.

[6] B. Csákány, Minimal clones – a minicourse, *Algebra Universalis* **54** (2005) 73–89.

[7] O. Ekin, S. Foldes, P. L. Hammer, L. Hellerstein, Equational characterizations of Boolean function classes, *Discrete Math.* **211** (2000) 27–51.

[8] A. Feigelson, L. Hellerstein, The forbidden projections of unate functions, *Discrete Appl. Math.* **77** (1997) 221–236.

[9] E. Fried, H. K. Kaiser, L. Márki, An elementary approach to polynomial interpolation in universal algebras, *Algebra Universalis* **15** (1982) 40–57.

[10] E. Lehtonen, Descending chains and antichains of the unary, linear, and monotone subfunction relations, *Order* **23** (2006) 129–142.

[11] E. Lehtonen, Á. Szendrei, Equivalence of operations with respect to discriminator clones, *Discrete Math.* **309** (2009) 673–685.

[12] N. Pippenger, Galois theory for minors of finite functions, *Discrete Math.* **254** (2002) 405–419.

[13] M. Pouzet, I. G. Rosenberg, Small clones and the projection property, arXiv:0705.1519v1.

[14] M. Pouzet, I. G. Rosenberg, M. G. Stone, A projection property, *Algebra Universalis* **36** (1996) 159–184.

[15] R. W. Quackenbush, A survey of minimal clones, *Aequationes Math.* **50** (1995) 3–16.

[16] I. G. Rosenberg, Minimal clones I: the five types, *Lectures in Universal Algebra (Proc. Conf. Szeged 1983), Colloq. Math. Soc. János Bolyai* **43**, North-Holland, 1986, pp. 405–427.

[17] A. Salomaa, On essential variables of functions, especially in the algebra of logic, *Ann. Acad. Sci. Fenn. Ser. A I. Math.* **339** (1963) 3–11.

[18] S. Świerczkowski, Algebras which are independently generated by every *n* elements, *Fund. Math.* **49** (1960) 93–104.

[19] C. Wang, Boolean minors, *Discrete Math.* **141** (1991) 237–258.

[20] R. Willard, Essential arities of term operations in finite algebras, *Discrete Math.* **149** (1996) 239–259.

[21] I. E. Zverovich, Characterizations of closed classes of Boolean functions in terms of forbidden subfunctions and Post classes, *Discrete Appl. Math.* **149** (2005) 200–218.

(M. Couceiro) Department of Mathematics, University of Luxembourg, 162a, avenue de la Faïencerie, L–1511 Luxembourg

E-mail address: `miguel.couceiro@uni.lu`

(E. Lehtonen) Department of Combinatorics and Optimization, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, N2L 3G1, Canada

Department of Mathematics, Tampere University of Technology, P.O. Box 553, FI–33101 Tampere, Finland

E-mail address: `erkko.lehtonen@uni.lu`